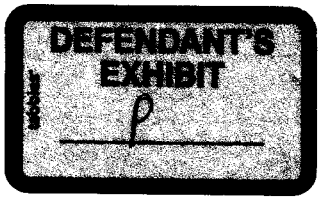
UNITED STATES DISTRICT COURT
MIDDLE DISTRICT OF NORTH CAROLINA
File No. 1:12-CV-1349



UNITED STATES OF AMERICA, Plaintiff,	)	
v.	)	EXPERT WITNESS REPORT (David L. Banks)
TERRY S. JOHNSON, in his official capacity as Alamance County Sheriff, Defendant.	)	

Report on Statistical Analyses in re *United States v. Terry S. Johnson*

David Banks, Department of Statistical Science, Duke University

1 Introduction

This report describes all of the statistical analyses that were done to assess whether or not there is numerical evidence that the Alamance County Sheriff's Office did, or did not, target Latinos for arrest. I was retained by the Turrentine Law Firm, PLLC, and my charge was to provide an impartial review of the data. I was assisted in my analyses by Daniel Heard, a fourth-year Ph.D. student in the Department of Statistical Science at Duke University.

At an initial meeting with S. C. Kitchen, attorneys for Alamance County, and members of the Alamance County Sheriff's Office, it was determined that the alleged prejudicial arrests could have occurred in three ways:

- That the Alamance County Sheriff's Office selected checkpoints sites that were more likely to encounter Latino drivers.
- That checkpoint screening outcomes led to targeted arrests for Latino drivers.
- That Latinos were more likely to be stopped for minor violations in non-checkpoint situations.

Using data supplied from the Alamance County Sheriff's Office, the U.S. Census Bureau, the North Carolina state government, and additional sources, I examined each of these three possibilities.

Based on these statistical analyses, I conclude that:

- Regarding checkpoint siting, I am confident that there is no evidence of malfeasance.
- Regarding checkpoint scrutiny, I could find no evidence that directly addressed that question. But I am confident that checkpoint practices in Alamance County are statistically similar to the practices in several other counties in central North Carolina.

- Regard arrest versus citation outcomes for people stopped at checkpoints, I found no significant differences between Hispanics and non-Hispanics.

- Regarding non-checkpoint traffic stops, I am confident that there is no evidence that Hispanic drivers are more likely to be stopped than non-Hispanic drivers. Also, it is clear that Alamance County stops Hispanics at a lower rate than many other central North Carolina counties.

2 Checkpoint Siting

The checkpoint process in Alamance County proceeds as follows:

1. There is a decision to conduct a checkpoint. All checkpoints examine driver's licenses; additionally, the checkpoint may look at registration, test sobriety, or confirm seatbelt use.

2. The location is chosen, and must satisfy a number of criteria specified in official policy.

3. A checkpoint authorization form is completed and signed by the supervisor.

4. The checkpoint is conducted. Every vehicle is examined unless safety or traffic considerations require temporary suspension of the checkpoint.

5. If there is no violation of the kind targeted in the checkpoint, then the driver proceeds without an official record being created. But if a violation is seen, then the deputy treats it as if it were a traffic stop, and the outcomes may be an arrest, a citation, a written warning, or an oral warning. Arrests, citations, and written warnings create an official record, but an oral warning does not.

In this framework, it would be possible for Hispanics to be improperly targeted either through the selection of checkpoint locations or through their treatment when stopped at a checkpoint.

In terms of checkpoint siting, there is no inventory of all possible sites in Alamance County that satisfy the policy constraints. For example, there must be room to pull over a reasonably large number of cars, there must be clear lines of sight for chase cars to pursue checkpoint evaders, and

the traffic flow cannot be so great that the checkpoint imposes serious delay, nor so low that the effort is unproductive.

Instead of an inventory, the checkpoint authorization forms provided a list of all 305 sites that were used between 2009 and 2012, inclusive (and I was told that records of locations from previous years were not available). Many of the 305 sites were used more than once. I treated this list as if it were a inventory of all feasible checkpoint sites, and, in my judgment, this assumption is a reasonable approximation.

I also obtained from Alamance County attorneys a list of the three locations that the Department of Justice identified in their initial complaint as targeted Latino communities in Alamance County (i.e., Seamster's Mobile Home Park, Rocky Top Mobile Home Park, and Calloway (Creekwood) Drive Mobile Home Park). Later, at the request of Alamance County attorneys, I added three additional mobile parks (Clover Creek, Oliver I, and Oliver II), since these had been subsequently identified by the Department of Justice as areas of concern.

The strategy for the statistical analysis was to determine whether the Alamance County Sheriff's Office (ACSO) had placed checkpoints near one or more of the Latino communities with greater probability than would have occurred by chance. If so, this could be construed as evidence of impermissible profiling.

To perform the analysis I used a permutation test. First, Daniel Heard and I calculated the shortest distance (along the roadway) from each of the checkpoint sites to the nearest Latino community. Then we ran 1000 simulations. Each simulation picked n_i checkpoint sites at random, where n_i is the number of checkpoints set up in year i, for $i = 2009, 2010, 2011, 2012$. The random selection was set up so that all checkpoint locations had equal probability of being chosen. It was possible for the same site to be chosen more than once in a year (as happens in practice).

For each simulation, the program calculated the sum of the shortest distances from each chosen checkpoint to the nearest Latino community. If the sum of the shortest distances from the checkpoints selected by the ACSO to the nearest Latino community tended to be smaller than the sums produced by the simulation, this would suggest that checkpoint sites had been chosen that were improbably near to Latino communities. In statistical parlance, the null hypothesis was that the ACSO had selected sites equiprobably from among the available sites, and the alternative hypothesis was that it had selected sites that were near Latino communities. In legal terms, the

null hypothesis is that ACSO did not target Latinos, and the alternative hypothesis is that it did.

Based on two sets of 1000 simulations, I found the P-values for the permutation tests using first the initial three locations identified by the Department of Justice, and then the expanded set of all six locations its lawyers had flagged. These values ranged between 0.21 and 0.47, as shown in the following table. (Note: A P-value is the probability of obtaining data that support the alternative hypothesis as or more strongly than the data observed, when the null hypothesis is correct.)

Table 1: The P-values, or significance probabilities, obtained from a permutation test of the null hypothesis that checkpoint sites were chosen at random from among the 305 historical sites in Alamance County. There are separate tests for each of the four years, and there are two alternative hypotheses: that sites are improbably near a set of three Latino communities initially identified by the Department of Justice, and that sites are improbably near an enlarged set of six communities subsequently identified by the Department of Justice.

	Year			
P-value	2009	2010	2011	2012
three communities	0.23	0.34	0.33	0.43
six communities	0.21	0.36	0.32	0.47

To combine the evidence from the P-values across all four years, I used Fisher's method:

$$-2 \sum_{i=2009}^{2012} \ln(P_i) \sim \chi^2_8.$$

Under the null hypothesis of random siting, the test statistic (i.e., the right-hand side) has a chi-squared distribution with degrees of freedom equal to twice the number of P-values; under the alternative hypothesis that sites are close to Latino communities, the test statistic will tend to be larger.

When three communities are used, the test statistic is 9.00 and the P-value is 0.34. When all six communities are used, the test statistic is 8.95, so the pooled P-value is 0.35. The statistical convention is that only P-values less than 0.05 are interesting, and that strong support for the alternative hypothesis does not arise unless the P-value is less than 0.01.

In summary, if the ACSO had chosen among the Alamance County checkpoints with equal probability, then between 21% of the time and 47% of the time it would have selected sites that

were as close or closer to Latino communities than the locations that were actually used. So there is no statistical evidence that checkpoints were sited in ways that targeted the Latino population.

2.1 Data

The list of unique checkpoint sites was provided by the ACSO to David Banks in the form of an Excel spreadsheet. The initial list of three Latino communities and their locations was sent to me on June 21, and the list of additional Latino communities and their locations was provided on July 31.

2.2 Assumptions

All statistical analyses make assumptions. I believe this analysis made plausible and impartial assumptions, and that nearly all statisticians and law enforcement experts would agree with these.

The most critical assumption in the permutation test is that the list of checkpoint sites is representative of the feasible checkpoint sites in Alamance County. If the ACSO had chosen a large proportion of the locations so as to be near a Latino community, then the permutation test would overestimate the P-value, as compared to a permutation test based upon all feasible sites. But even in that extreme case, the observed P-values indicate that the checkpoints used were much less targeted than such aggressive site selection would have allowed, and thus are inconsistent with a theory of ethnic profiling.

The second, and more minor, assumption is that checkpoint locations were chosen independently from the list of 305 sites. Independence implies that the site chosen today has nothing to do with the site chosen yesterday or the day before. As a matter of administrative practice this seems unlikely, but the impact of small dependencies over the course of a year, or four years, will be very small and should have no effect upon the conclusion.

3 Checkpoint Screening Outcomes

A second way in which Latinos could be targeted would be for them to receive greater scrutiny at checkpoints than non-Latino groups, or for them to have worse outcomes at checkpoints than

non-Latino groups.

3.1 Increased Scrutiny

In North Carolina, all checkpoint stops entail examination of driver's licenses. And, during the time period indicated in the Department of Justice complaint, no undocumented person could legally obtain a North Carolina driver's license. So one expects, a priori, that Latino drivers are more likely to be arrested or cited at a traffic checkpoint.

There are no data available which directly speaks to screening intensity at checkpoints. However, I did obtain data on the proportion of Hispanics who were charged as part of a checkpoint stop, and I note that, from the 2010 U.S. Census, 11.6% of Alamance county is Hispanic, but Hispanics account for 36.68% of the people charged at a checkpoint stop.

This is a large difference: the arrest/citation rate for Hispanics is 3.16 times their representation in the population. But there are at least three legally permissible reasons that could account for this discrepancy:

1. People without drivers' licenses are charged at checkpoint stops that examine licenses and registration. Since, during the time span covered in this complaint, undocumented people could not obtain a license, this will lead to excess Latino arrests.

2. It may be that Latinos tend to drive on roads where checkpoints are more commonly sited, or that Latinos tend to drive at times of the day when checkpoints are more likely to be set up.

3. It may be that Latinos are more likely to commit certain kinds of offenses which would lead to arrest or citation at a checkpoint.

It is possible that all three of these reasons are jointly responsible for the discrepancy in checkpoint rates of arrest or citation between Hispanics and non-Hispanics.

The Foundation for Traffic Safety (2011) reports that unlicensed drivers are more dangerous, and are more likely to flee the scene of an accident, so it is presumably part of the law enforcement mission to screen for licensed drivers. Tautologically, such screening will have disparate impact

on the undocumented community, and necessarily must increase the relative rate of Hispanics who are charged at traffic checkpoints.

Regarding the possible tendency for Hispanics to drive on roads that are more likely to have checkpoints, or to drive at times of the day when checkpoints are more likely to be set up, we have no data that can estimate the magnitude of this effect, and we see no practical way to acquire such data. But we raise this point because it indicates the kind of information gaps that arise in practice, and the consequent uncertainty in distinguishing intentional ethnic targeting from spurious associations caused by confounding factors.

Regarding the possibility that Latinos are more likely to commit certain kinds of offenses that could lead to arrest or citation at checkpoints, there are several reasons to think this is plausible. First, illegal immigrants are disproportionately male and disproportionately young (cf. Carriquiry et al., ch. 3, 2013). If their driving practices are like those of young Anglo males, this means they are more likely to fail sobriety checkpoints that screen for drugs and alcohol (Zador, Krawchuk, and Voas, 2000).

Stronger evidence of difference, by legal status, in the chance of failing a sobriety test is given by Guterbock et al. (2010). This study of the implementation of the 287(g) program in Prince William County, Virginia reports in Table 8.2 of Chapter 8 that 13% of DUI arrests were illegal immigrants. But chapter 7, pp. 76-77, claims that the conservative estimate is that there were at most 16,200 undocumented residents in Prince William County in 2008, and the American Community Survey puts the county population at 364,734, so the percentage of illegal immigrants is not greater than 4.4% (so the ratio of arrests to population is 2.95). This disparity is not an isolated case—the National Highway Transportation Safety Administration (2010) reports that alcohol abuse among Latino youths is greater than in the non-Hispanic community, and it undertook three studies to improve public health messaging to this demographic. One of the study sites was Durham, NC, for which the Latino population may be similar to that in Alamance County. In particular, the report indicates that in 2001, 29% of all traffic crashes in Durham County involved Latino drivers, although Latinos comprise only 9% of the population (a ratio of 3.22).

The preceding discussion cannot provide strong statistical evidence one way or the other regarding the possibility that the Hispanic community may have received extra scrutiny at check-

points in Alamance County. We know of no data set which could speak to this. However, it is possible to compare the rate at which Hispanics are arrested or cited at traffic points in Alamance County to similar rates in other central North Carolina counties. Figure 1 provides that comparison in terms of the disparity rate, where the disparity rate is defined as the ratio of the proportion of checkpoint stops involving Hispanics to the proportion of the county population classified by the U.S. Census as Hispanic. By this criterion, if the ACSO is targeting the Hispanic population at checkpoints, it is doing less of this than three other North Carolina Counties, and is comparable to at least three others among the thirteen central counties.

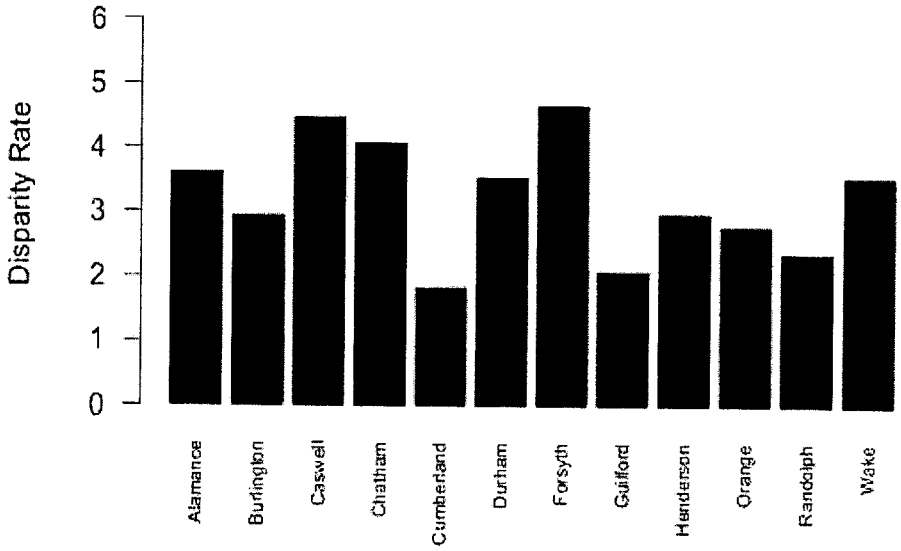


Figure 1: Bar chart of the ratio of Hispanic checkpoint stops to the Hispanic population, by county. Values greater than 1 indicate that more Hispanics are arrested or receive citations at traffic stops than would be expected if violations were proportional to representation in the community.

It should be emphasized that the county-level data used in this section refer to checkpoints set up by sheriff's offices in these counties, and do not include stops by city officers in those counties.

3.2 Arrests and Citations

As an additional investigation, I compared the arrest rates and citation rates for Hispanic and non-Hispanic drivers who were stopped at checkpoints. If the ACSO were targeting Hispanics, this could be reflected in higher rates of arrest, as compared to citation.

From 2009 through 2012, the ACSO conducted 453 checkpoints, based upon the checkpoint authorization forms. On my instructions, Daniel Heard drew a random sample of 45 ($\approx$ 10%) of these checkpoints, and hand-matched the arrest or citation form to the traffic stop report. For each of these stops, he also recorded the ethnicity of the driver and any other passengers in the vehicle.

For all checkpoint stops in the analysis, there were 21 arrests in 110 stops, giving an arrest rate of 19.1%. A 95% confidence interval for the checkpoint arrest rate is 12.5% to 27.4%. For checkpoint stops involving Hispanics, there were 6 arrests in 40 stops, giving an arrest rate of 15%. A 95% confidence interval for the checkpoint arrest rate for Hispanics is 3.3% to 24.6%. For checkpoint stops involving non-Hispanics, there were 15 arrests in 70 stops, giving an arrest rate of 21.4% with a 95% confidence interval of 11.8% to 31.0%. Clearly, non-Hispanics were arrested with greater frequency than Hispanics, but the confidence intervals show considerable overlap and the difference is not statistically significant (the P-value is 0.20 for a one-sided test of the alternative hypothesis that Hispanics are more likely to be arrested).

I also performed an analysis that restricted the scope to checkpoint stops for No Operator's License (NOL) violations. Among all NOL checkpoint stops, there were 2 arrests in 47 stops, giving an arrest rate of 4.3%. A 95% confidence interval for the NOL checkpoint arrest rate is 0.74% to 15.7%. For NOL stops involving Hispanic drivers, there were 2 arrests in 33 stops, giving an arrest rate of 6.1%. A 95% confidence interval for the NOL checkpoint arrest rate for Hispanics is 1.7% to 21.6%. For NOL stops involving non-Hispanic drivers, there were 0 arrests in 14 stops. Using the Clopper-Pearson method, necessary because of the 0 count for non-Hispanics, the exact 95% confidence interval is 0% to 23.1%. Thus, non-Hispanics are arrested with less frequency than Hispanics, but the confidence intervals overlap and the difference is not statistically significant (the P-value is 0.17).

3.3 Data

Regarding the possibility of increased scrutiny at checkpoints, we had little direct data. The main information was the combined arrest and citation counts at traffic checkpoints for Hispanics (`http://trafficstops.ncdoj.gov`) and the size of the Hispanic populations in various North Carolina counties (the U.S. Census Bureau). The estimates of the size of the Hispanic populations were based on the unadjusted Census figures, to better capture non-residents.

Alamance County was not atypical, but we note that there is large variation in the reported numbers which, to a statistician, suggests that unknown factors are at play.

Also, as a minor point, we are skeptical of the numbers reported by Guilford County. They indicate that they set up only 29 checkpoints during a three-year period, a number so small that it is difficult to credit.

Regarding checkpoint outcomes, the counts on the citation and arrest rates at checkpoints in Alamance County were obtained by Daniel Heard, by hand-matching arrest reports and citation reports to traffic stop reports. These reports are on file at the ACSO. He took a random 10% sample of all checkpoints set up between 2009 and 2012.

3.4 Assumptions

The decision to use the unadjusted Census figures, rather than the adjusted Census figures, is a judgment call. Fridell (2004, chapter 5) argues that various adjustments are useful, to account for differences in automobile ownership and the recurrent difficulty of enumerating people in hard-to-reach demographic groups. But Fridell's analysis does not take account of the unique aspects of the Latino population. First, the Latino community tends to be undercounted, which means that the use of the unadjusted numbers is conservative—it will tend to make the counties look worse than they actually are. Second, the Latino community is thought to contain a number of non-residents who stay with friends or relatives for extended periods of time, and the standard adjustment would exclude these. Also, the adjustments should change the numbers in small ways, and I do not think it would substantially alter the conclusions.

If the 3.16 discrepancy factor is seen as evidence of profiling, then one must assume that Hispanic drivers are in all ways similar to non-Hispanic drivers (e.g., they have similar patterns

driver's license ownership, similar patterns of DUI driving, and they travel at similar times of day along similar roads, have cars that have similar rates of registration and comparable mechanical condition, and so forth). Several of the studies cited in this section indicate that such strong assumptions are unrealistic.

Nonetheless, this is a large discrepancy and one that is seen across many North Carolina counties. We encourage law enforcement authorities to collect data that would enable a better understanding of the underlying reasons. Without such data, we see no way to statistically distinguish profiling from proper conduct based on checkpoint citations/arrests.

4 Targeted Traffic Stops

The third way in which the ACSO might have discriminated against Latinos would be for deputies to be more aggressive in pulling over and investigating drivers who appear to be Latino. If Anglos who drive five miles over the speed limit are ignored but Latinos are not, then this would constitute targeting.

The statistical difficulty in addressing this question concerns the appropriate measure of exposure (or benchmark). If Latinos drive more than non-Latinos, then they are more likely to stopped by a deputy. If they tend to drive at times of the day when more patrol cars are on the road, or in areas that are more heavily patrolled, then they are more likely to be stopped. If their cars are older and thus more likely to have mechanical problems (e.g., a broken turn signal), then they are more likely to be stopped. If they have a larger proportion of males in the 18 to 25 age range than Anglos (and if Latino males have the same driving pattern as Anglo males), then Latinos will be stopped more often. Ideally, one would adjust for all these factors and more, but there is no reliable data for such adjustment—we cannot calculate a risk-weighted measure of exposure.

The difficulty in estimating exposure in the context of possible racial or ethnic profiling is considered at length in Fridell (2004). Her discussion of benchmarking considers use of adjusted Census data, DMV data, "blind" enforcement data such as stoplight cameras, visual identification of driver race or ethnicity from roadside surveys, and other techniques. As emphasized in that book, all of these procedures are problematic. Among them, the best is to use the ethnicity of the not-at-fault driver in accidents to infer the proportion of miles driven by that ethnic group

(fatal accident data are available at the U.S. Department of Transportation FARS database, and non-fatal accident data could, in principle, be recovered by hand review of accident reports in Alamance County).

We did not pursue either data set because, in the context of undocumented residents, the not-at-fault driver is incentivized to not report minor accidents (to avoid deportation). Reporting rates should be much more reliable for fatal accidents, but these numbers are so low in Alamance County that the resulting confidence intervals on estimates would be uselessly wide. Also, the not-at-fault method does not automatically adjust for such things as the mechanical condition of the car or differences in driving styles by age and gender.

Given these obstacles, I chose to do cross-county comparisons instead. North Carolina maintains records on traffic stops at `http://www.trafficstops.ncdoj.gov/`, and these data are broken out by gender and ethnicity. The data showed that in Alamance County there were 14481 non-checkpoint stops during 2009-2012, of which 1668 involved Hispanic drivers (of which Latinos are surely the vast majority). Thus 11.5% of the non-checkpoint stops (which entail more officer discretion than checkpoint stops), involved Hispanics. For comparison, the unadjusted Census figures show 11.6% of Alamance County is Hispanic. In this comparison I use the unadjusted Census figures, since they include non-resident Hispanics; for the Latino population, this could easily account for a few percentage points.

First, we note that the deficiency of 0.1% in non-checkpoint stops provides no evidence that ACSO deputies are targeting Hispanic drivers. On the face of it, if there is any effect, the deputies are favoring them. (Of course, it is also possible that undocumented drivers are risk averse, and thus more mindful of speed limits and use turn signals conscientiously, and thus are less likely to be stopped.)

Table 2 shows the non-checkpoint (NCP) stop percentages in 13 central North Carolina counties, along with the P-values for testing the statistical significance of differences between the NCP stop percentages and the unadjusted Census estimates of the percentages of the county populations that are Hispanic.

Examination of the P-values indicates that, if ethnic disparity in stop rates is a sign of profiling, then Alamance County is not targeting Hispanics. But other counties, with P-values of 0.01 or so, appear more problematic.

Table 2: The significance probabilities, or P-values, associated with 13 tests of the null hypothesis that the non-checkpoint stop rates for Hispanic drivers are equal to the proportion of Hispanic residents in 13 counties in central North Carolina. These are one-sided tests, so the alternative is that Hispanics drivers are stopped disproportionately often.

County	Hispanic NCP stop %	Hispanic Census %	P-value
Alamance	11.50%	11.6%	0.96
Cabarrus	5.41%	9.7%	0.99
Caswell	3.73%	3.2%	0.05
Chatham	12.16%	12.9%	0.98
Durham	10.45%	13.4%	0.99
Forsyth	6.32%	12.4%	0.99
Gaston	5.10%	6.1%	0.85
Guilford	8.29%	7.4%	0.01
Mecklenburg	14.54%	12.5%	0.02
Orange	18.04%	8.2%	<0.01
Randolph	21.34%	10.7%	<0.01
Rockingham	5.68%	5.7%	0.49
Wake	11.64%	10.0%	<0.01

Additionally, Fig. 2 displays the standardized disparity rate (SDR) of non-checkpoint stops involving Hispanics for Alamance and 12 other nearby counties. The SDR is defined as

$$\text{SDR} = \frac{\text{proportion of NCP Hispanic stops}}{\text{Hispanic proportion of the population}} - 1$$

so that values near zero indicate that stop rates match the population rates, values greater than zero indicate higher stop rates for Hispanics, and values less than zero indicate lower stop rates for Hispanics.

The main conclusion from Fig. 2 is that, compared to many of its neighbors, Alamance County stops a remarkably small number of Hispanic drivers.

As before, it should be emphasized that the county-level data used in Table 1 and Fig. 2 refer to checkpoints by sheriff's offices in these counties, and do not include stops by city officers in those counties.

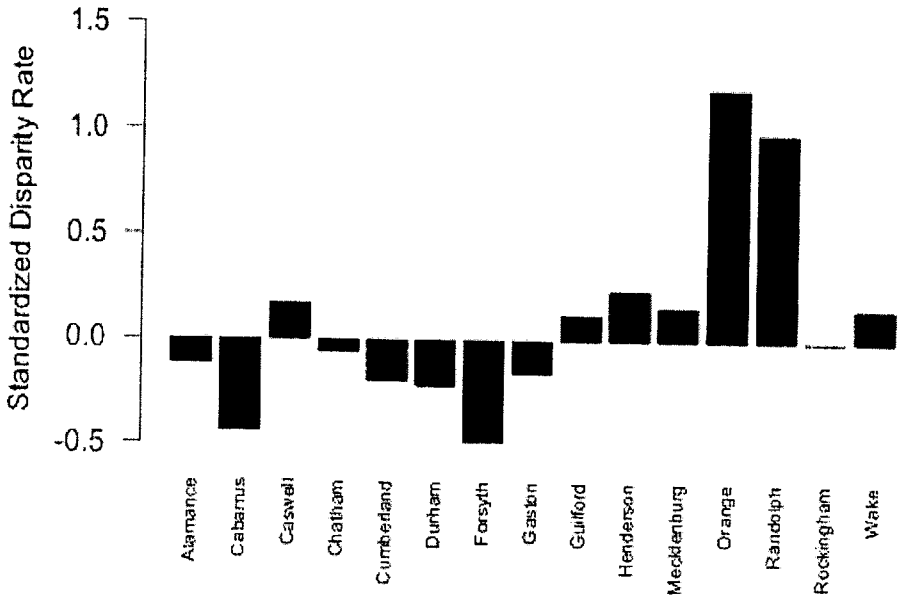


Figure 2: A bar chart of the standardized disparity rate for Hispanic non-checkpoint stops versus the Hispanic population, broken out by county. Values greater than zero imply more stops that would be expected if stops were proportional to representation in the community.

4.1 Data

The North Carolina government site `http://www.trafficstops.ncdoj.gov/` provided the data on traffic stops, broken out by race and ethnicity. Population estimates came from the U.S. Census.

4.2 Assumptions

The hypothesis tests assume that the traffic stops are made at random, in the sense that the probability of a stop, conditional on the circumstances, is constant. This is probably not strictly true; e.g., there may be a directive to crack down on speeding in school zones during the first month of the school year. But I believe, and think that nearly all statisticians would agree, that the impact on the conclusions from plausible levels of violation of the assumption is small.

As previously discussed, it is an open question whether or not one should use unadjusted

or adjusted Census estimates. I used the unadjusted estimates, since I believe it is a conservative choice. As before, I do not think that, for the Hispanic community, the small differences I anticipate from adjustment would alter the main conclusions.

5 Vehicle Miles Traveled

Besides examining the three ways in which Sheriff Johnson could have targeted Hispanics, we also considered a number of other analyses that might have provided data to inform those investigations. In particular, we discussed using Hispanic enrollment in schools to improve estimates of the undocumented population, using observational studies to estimate the proportion of vehicle miles traveled by people whose features appear Hispanic, and using data on commuting patterns to estimate the vehicle miles traveled by Hispanics in Alamance County.

Ultimately, we did no analyses with school enrollment data, and no observational study was undertaken (it seemed impractical, and we had no confidence that such a study could be valid or have generalizability). However, we did examine the commuting data, which broke out commuting miles driven by Hispanic and non-Hispanic drivers, and by Alamance County residents and non-residents.

The reason for examining the commuting data is twofold:

1. If the ratio of the number of commuting miles driven by Hispanic drivers to the total miles they drive is the same as the ratio of the number of commuting miles driven by non-Hispanics to the total miles driven by non-Hispanics, then this provides an estimate of relative vehicle miles traveled and thus a (non-risk-adjusted) measure of exposure.

2. Alamance County contains major highways that are used by non-residents for commuting. Some of these drivers are arrested or receive citations, but their presence in the data is not properly accounted for by Census estimates of residents. Use of non-resident commuting data could improve that aspect of the analysis.

For these reasons, we briefly studied the commuting data. Ultimately, it did not directly impact our analyses of the three possible ways in which Latinos could have been profiled. We chose not to use it because it entailed assumptions that were difficult to verify, such as similar ratios of

commuting to non-commuting driving, and equivalence of stop rates on highways to stop rates on secondary roads.

Nonetheless, the main results from that exploration were that 60.3% of vehicle miles traveled within Alamance County are by non-Hispanic residents of Alamance County, 29.76% of vehicle miles traveled within Alamance County are by non-Hispanic residents of other counties, 6.43% of vehicle miles traveled within Alamance County are by Hispanic residents of Alamance County, and 3.51% of vehicle miles traveled within Alamance County are by Hispanic residents of other counties. Thus, this method estimates that 9.94% of the vehicle miles traveled in Alamance County are driven by Hispanics.

If this estimate is correct, it would have no effect on the checkpoint siting analysis. For the analysis of checkpoint screening intensity, it would increase the effective arrest/citation rate for Hispanics from 3.16 to 3.69, but one would have to believe that checkpoints are sited on major commuting highways at the same rate that they are placed on secondary roads. For the analysis of traffic stops, it would change the SDR from -0.01 to 0.16 (again, one would have to assume that stop rates on highways are like stop rates on other roads).

5.1 Data

The data on commuting patterns was gathered by the U.S. Bureau of the Census and is at `http://lehd.did.census.gov/`.

5.2 Assumptions

Using the commuting data requires strong assumptions that are difficult to verify. After considerable discussion, it remained unclear whether trying to use (potentially unreliable) self-reports of commuting miles to infer total driving would improve any of the analyses I performed. In particular, there are issues of ethnic differences in ride-sharing, exposure avoidance by undocumented drivers, non-comparable law enforcement practices on commuting highways, and so forth.

6 Conclusions

This study was undertaken to examine the extent of statistical information as to whether or not the Alamance County Sheriff's Office may have targeted Hispanics for arrest. We could imagine only three ways in which such targeting might be implemented: checkpoint siting, checkpoint outcomes, and non-checkpoint stops.

Regarding checkpoint siting, I am confident that there is no evidence of malfeasance.

Regarding checkpoint scrutiny, I was not able to find evidence that could directly address this question. But I am confident that the checkpoint practices in Alamance County are statistically similar to the practices in many other counties in central North Carolina.

Regarding arrest versus citation outcomes for people stopped at checkpoints, I found no significant differences between Hispanics and non-Hispanics.

Regarding non-checkpoint traffic stops, I am confident that there is no evidence that Hispanic drivers are more likely to be stopped than non-Hispanic drivers. Also, I am confident that Alamance County stops Hispanics at a lower rate than many other central North Carolina counties.

7 References

Carriquiry, A., Banks, D., Brownell, P., Fienberg, S. E., Handcock, M., Hanson, G., Lesser, V., Orrenius, P., Passel, J., Riosmena, F., and Saucedo, S. (2013). *Options for Estimating Illegal Entries at the U.S.-Mexico Border*, National Research Council, Washington, D.C.

Foundation for Traffic Safety (2011). *Unlicensed to Kill*, American Automobile Association, Washington, D.C.

Fridell, L. (2004). *By the Numbers: A Guide for Analyzing Race Data from Vehicle Stops*, Police Executive Research Forum, Washington, D.C.

Guterbock, T. M., Koper, C. S., Vickerman, M., Taylor, B., Walker, K. E., and Carter, T. (2010). *Evaluation Study of Prince William County's Illegal Immigration Enforcement Policy*, Center for Survey Research, University of Virginia.

National Highway Transportation Safety Administration (2010). *Promising Practices for Addressing Alcohol-Impaired Driving within Latino Populations: A NHTSA Demonstration Project*, Department of Transportation, Washington, D.C.

Zador, P., Krawchuk, S., and Voas, R. (2000). "Alcohol-Related Relative Risk of Driver Fatalities and Driver Involvement in Fatal Crashes in Relation to Driver Age and Gender: An Update Using 1996 Data," *Journal of Studies on Alcohol and Drugs*, 61, 387–395.

This report was prepared by David L. Banks.